REVIEW

Does methylphenidate improve academic performance? A systematic review and meta-analysis

Anne Fleur Kortekaas-Rijlaarsdam[1] · Marjolein Luman[1] · Edmund Sonuga-Barke[2] · Jaap Oosterlaan[1]

Abstract
Academic improvement is amongst the most common treatment targets when prescribing stimulants to children with ADHD. Previous reviews on stimulant-related academic improvements are inconclusive and focus on task engagement. Recent literature suggests outcome-domain-specific medication effects that are larger for productivity than for accuracy. The aims of this study are quantifying methylphenidate effects on academic productivity and accuracy for math, reading, spelling; exploring the mediating or moderating effects of symptom improvements, demographic-, design- and disorder-related variables. PubMed, EMBASE, ERIC and PsycINFO were searched for articles reporting methylphenidate effects on academic productivity and accuracy. Thirty-four studies met entry criteria. Methylphenidate improved math productivity (7.8% increase, $p < .001$); math accuracy (3.0% increase, $p = .001$); increased reading speed (SMD .47, $p < .001$) but not reading accuracy. None of the mediators or moderators tested affected methylphenidate efficacy. Academic improvements were small compared to symptom improvements; qualitative changes limited to math. Clinicians should take this discrepancy into account when prescribing medication for ADHD.

Keywords ADHD · Methylphenidate · Academic · Math · Reading · Meta-analysis

Introduction

The neuro-developmental disorder Attention-Deficit Hyperactivity Disorder (ADHD) is characterized by inattention, hyperactivity and impulsivity [1]. These symptoms are associated with academic problems such as lower grades, grade repetition and increased school drop-out [2–5]. Academic improvement is a common treatment target for children with ADHD with stimulant medication being the most commonly prescribed treatment [6]. Stimulants are clinically effective in reducing ADHD symptoms in the short- and medium term

✉ Anne Fleur Kortekaas-Rijlaarsdam
  A.F.Kortekaas-Rijlaarsdam@vu.nl

[1] Clinical Neuropsychology section, Vrije Universiteit Amsterdam, Van der Boechorststraat 1, 1081 BT Amsterdam, The Netherlands

[2] Department of Child and Adolescent Psychiatry, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, England, UK

[7, 8]. Moreover, there is evidence that its benefits extend to improvements in cognition relevant for academic performance [9].

In the past, there have been several reviews of the effects of stimulant medication on academic performance [10–16]. These reviews report little evidence for positive effects [10, 11, 16, 17]. However, the first meta-analysis [13], demonstrated 9.7–14.4% ($p < .001$) improvements with stimulant medication compared to placebo in seatwork productivity (number of assignments completed) and on-task behavior (amount of time actively spend on seatwork). However, the effect of stimulant medication on the quality of academic performance (academic accuracy) was less clear-cut: in one analysis, only a third of the studies reported effects of medication on academic accuracy and the pooled effect was not significant [13].

The existing reviews [10–16] and meta-analysis by Prasad et al. [13] have a number of limitations. First, because the negative association between ADHD symptoms and both reading and math is stronger than the association between ADHD symptoms and spelling [5] and there is evidence from recent medication trials that medication efficacy differs between academic subjects [18, 19], it is important to

independently assess the effects of medication on different academic subjects. The meta-analytic results of Prasad and colleagues [13] only applied to seatwork assigned by participant's teachers (independent of the academic subject) and, therefore, tell us little about the possible differential effects of stimulants on specific academic subjects. Second, Prasad and colleagues reported on the number of items which are correct, a measure that might be confounded by the number of items attempted (task productivity), rather than percentage correct. Therefore, improvements in items completed correctly may merely reflect increased productivity. It is thus important to distinguish between improvements in accuracy and productivity, especially because long-term studies suggest that improvements in test scores with medication are often not accompanied by improvements in longer term academic outcomes, such as grades and grade repetition [14, 15].

Ultimately, prior reviews have not resolved the key question of whether there are improvements in core academic skills or just improvements in academic productivity. Here we conducted a meta-analysis to address this issue. Our meta-analysis improved upon previous work in several aspects. First, in contrast to the previous reviews, we quantified both accuracy and productivity while distinguishing between the core academic subjects (math, reading and spelling). Second, we added 6 years of literature to the prior meta-analysis [13], which more than doubled the number of studies included. Third, these more recent studies allowed for exploration of the moderating effects of demographic and disorder-related variables (age, gender, ADHD subtype and severity, and commonly reported comorbidity with oppositional defiant disorder, conduct disorder and learning disorders) and study characteristics [medication release system, dosage, titration method, time of measurement (i.e., hours after intake) and trial duration] on medication efficacy. Fourth, we followed up on recent studies suggesting that academic improvements due to stimulant medication were partly mediated by behavioral improvements [19, 20]. Therefore, the current meta-analysis included symptom improvements and on-task behavior (% time on task) as potential mediators in the analyses.

Methods

This systematic review conformed to PRISMA [21].

Study selection

We included studies published in the English language in peer-reviewed journals that (1) evaluated the effects of stimulant medication on academic functioning; (2) included mostly (at least 80% of the sample) primary school children

(male and female) with a primary diagnosis of ADHD [established using the DSM-III, DSM-III-Revised (DSM-III-R) or DSM-IV/DSM-IV-Text Revisions (DSM-IV-TR) or ICD-10 criteria]; (3) evaluated the effects of methylphenidate (MPH) (immediate or extended release formulations or transdermal) on standardized achievement tests for math, reading or spelling; and (4) used a placebo-controlled crossover design or between-subject design. This review focused on primary school children as it is at this age that school teachers, concerned about academic performance, often drive the referral process, advising parents to seek help for their children's ADHD. Further, there is a big difference in medication use in primary school age students and high school students (e.g., non-compliance rates are much higher in high school, see for example [22]). The computerized databases PubMed, EMBASE, ERIC and PsycINFO were used to identify relevant studies up to October 2017. The following search terms and all possible equivalents were used to search article title and abstract: (1) disorder terms, e.g., 'ADHD'; (2) treatment terms, e.g., 'methylphenidate', including all brand names; (3) outcome terms, e.g., 'academic', 'school', 'classroom', 'math', 'reading', 'spelling', 'writing', 'on-task', 'off-task'. In case of missing or incomplete data, authors were contacted twice for additional data. When data were presented in graphs only, we used GetData Graph Digitizer version 2.26 [23] to extract the exact numbers, which was done successfully for one study [24]. In cases where multiple articles were based on the same sample, we selected the original, most comprehensive report on that study, resulting in the exclusion of four studies [25–28] as these data were originally described elsewhere [18, 29, 30]. See Fig. 1 for a flow diagram of the meta-analytic search and study selection. The first author (AK) and a second independent investigator reviewed titles and abstract for eligibility. Full texts were also reviewed by the first author as well as by an independent investigator. A third independent investigator conciliated discrepancies. Reference lists of included articles were searched for additional articles meeting the inclusion criteria.

A total of 3084 records were identified corresponding to 2594 unique articles. Thirty-four articles met the inclusion criteria for meta-analysis (Fig. 1). Study characteristics, including design, medication titration, dependent variables, mediators and moderators obtained from each study are displayed in supplementary material Table E1.

Measures and data extraction

Table E2 in supplementary material gives an overview of tasks and questionnaires used to assess academic outcomes, an overview of the selected mediator and moderator variables, as well as the measures derived from the academic tasks.
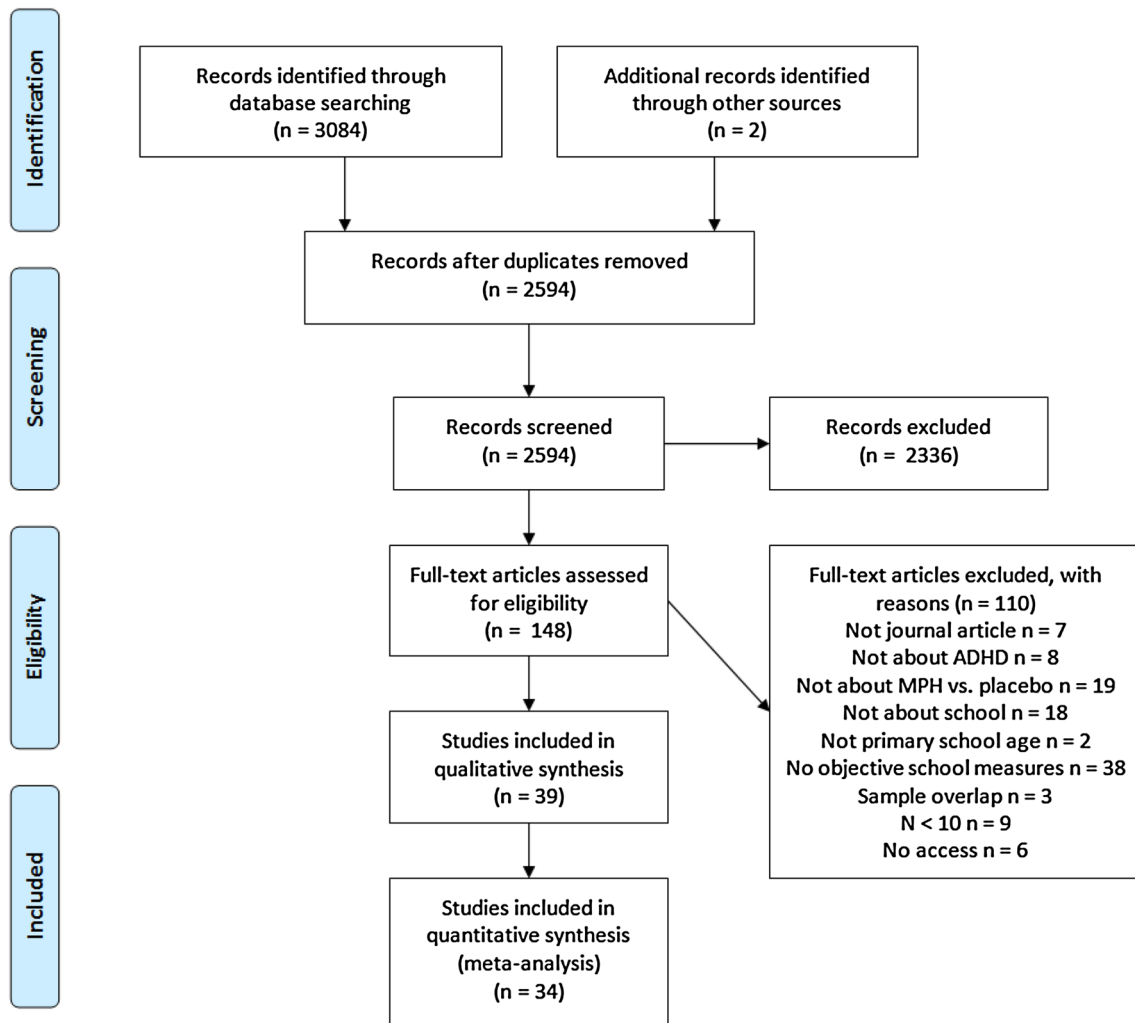
Fig. 1 Flow diagram of the meta-analytic search and study selection

Academic outcomes

Articles were included if they provided information about either accuracy or productivity scores for math, reading or spelling, or a combination of these. When mean accuracy and productivity scores were not reported as dependent variables, they were calculated by hand. Accuracy was calculated by dividing the mean number correct responses by the mean number items completed. Productivity was calculated by dividing the mean number of items completed by the total number of items.

Math and reading tasks were always speeded tasks requiring participants to complete as many items in a limited amount of time. Math tasks always consisted of simple math problems (addition, subtraction, multiplication and division) generally presented in ascending order of difficulty over a fixed period of time. Reading tasks consisted of a short passage text followed by multiple choice questions. Reading paragraphs were adapted to the student's reading level. Meta-analyses were performed for math accuracy, math productivity, reading accuracy and reading number attempted. The latter was chosen as an outcome because reading productivity could not be calculated as the total number of reading items differed per study and was not reported in combination with reading number attempted. Reading number attempted is an informative measure as the included studies used identical tasks and time limits. Spelling was measured by spelling lists assigned by teachers or taken from local school district lists. Only two out of three studies from our search reported spelling accuracy [31, 32], the third study only reported standardized means related to baseline scores [33]. As only two studies met inclusion criteria and minimum number of studies to perform a meta-analysis is three [34] we limited our analysis to a narrative description and qualitative synthesis.

Mediators

ADHD symptom improvements were included as mediators if means and standard deviations were available. Because the number of studies reporting on parent-rated symptom improvements or on-task behavior were limited ($n = 2$ and $n = 8$, respectively) and at least ten studies are recommended for reliable meta-regression [34], we only included teacher-rated symptom improvements in our mediator analysis and performed meta-regression for math accuracy ($n = 17$) and math productivity ($n = 11$). Teacher-rated symptom improvements were measured with standardized questionnaires, which were either derivatives from the Conners Rating Scale [35], the Strength and Weakness of ADHD symptoms and Normal Behavior (SWAN) rating scale [36, 37] or the Swanson, Kotkin, Agler, M-Flynn, and Pelham (SKAMP) rating scale. SKAMP ratings show high correlations ($r = .50$–$.84$) with Conners ratings scales [38]. Supplementary material Table E2 gives an overview of all questionnaires used in the studies included in this meta-analysis. Reliability and validity of all questionnaires used have been established [38, 39]. Scores were standardized (mean difference between conditions divided by SD of the placebo condition) for inclusion in the meta-regression. Mediators were investigated using meta-regression, using difference scores (MPH minus placebo).

Moderators

Because the number of studies was limited and at least ten studies are recommended for reliable meta-regression [34], we performed meta-regression analyses only for math outcomes. For math accuracy we tested the following moderators: (1) demographic moderators age (year); gender (percent male); (2) disorder-related moderators percent children diagnosed with ADHD-inattentive subtype; percent diagnosed with comorbid ODD or CD; parent-rated ADHD severity [standardized (mean divided by SD) baseline ADHD symptom ratings on standardized questionnaires, for an overview of questionnaires used to assess severity of ADHD symptoms, see supplementary material Table E2], and (3) study characteristics including release system (immediate versus extended release, the latter including transdermal); duration of the study conditions (days); time of measurement (post-dose, in hours); medication dosage (mg); titration method (clinical titration versus fixed dosages). For math productivity we tested demographic and disorder-related moderators: age, gender, percent diagnosed with ADHD-inattentive subtype, and study characteristics: release system, trial duration, and titration method. Insufficient number of studies reported on comorbid learning disorders. In case of doubt, authors were contacted.

Moderators were explored using meta-regression between the study samples' effect sizes for academic performance and the selected moderators. Mediator and moderator effects were studied separately for math accuracy and math productivity using meta-regression with a random model (method of moments) [34].

Statistical analysis

Statistical analyses were performed using SPSS version 21.0 [40] and Comprehensive Meta-Analysis software V3.0 [41]. Because accuracy and productivity measures are proportional measures that require effect sizes for binary data [34], risk differences (MPH minus placebo) were calculated. The standard errors of the risk difference were calculated [42] because included articles commonly reported on the number correct and the number completed and, therefore, the reported p values and standard deviations were not applicable to the calculated risk differences. Effect sizes were calculated for math accuracy, math productivity, reading accuracy and reading number attempted. In supplementary material Table E1, we provide a narrative description of the studies reporting on spelling. The derived effect sizes were weighted by their inverse variance to account for differences in sample size and error of measurement [34]. As heterogeneity may have been introduced using data from studies with different designs (i.e., differences in treatment duration, dosages) and different participants (e.g., differences in comorbidity), all meta-analytic effect sizes were calculated using a random effects model. The I^2 statistic was used to assess heterogeneity of effect sizes, where values of 25, 50 and 75% indicate low, moderate and high heterogeneity, respectively [43].

Rosenthal's fail-safe n was calculated to determine the number of studies with a null effect necessary to cancel out significant effect sizes, where fail-safe n values $> 5k + 10$ were considered robust and k refers to the number of samples on which the relevant effect size was calculated [44]. Further, Egger funnel plot asymmetry was used to assess publication bias [45]. Associations between effect size and sample size were investigated to assess the possibility that studies with small samples and large effect sizes were more easily published than studies reporting non-significant findings. All tests of significance were two sided with $\alpha = .05$. Risk of bias was estimated for each study based on Cochrane guidelines [46].

Results

A combined total of 1777 children from 34 different studies were included in the meta-analyses. Another 425 children from seven studies were included in the qualitative synthesis because results were either reported in figures only

or exact values were not reported, including six studies on math performance and three studies on spelling accuracy. Table E1 provides an overview of which studies qualified for meta-analysis and gives a narrative description of the results of those studies that did not qualify. Meta-analyses were conducted for math accuracy (29 studies, $N = 1528$) and math productivity (17 studies, $N = 912$). For reading, meta-analyses were conducted for reading accuracy (nine studies, $N = 207$) and number of items attempted (five studies, N = 100). Most studies (88.2%) used a placebo-controlled crossover design. Four studies (11.8%) used a between-subject design. In 73.5% of the studies, medication dosage was clinically titrated on symptom improvement before start of the trial. In the other 26.5% of the studies, dosages were fixed. When multiple dosages were used in randomized order, we included results from the dosage showing greatest effects on academic outcomes to optimize MPH efficacy (please see Table E1 for details). While all studies predominantly involved primary school children, one study also included children from middle school (aged 12–16, 16.5%).

Effects of MPH on academic performance

Table 1 provides an overview of all meta-analytic results, heterogeneity statistics and the results of the publication bias analyses.

Math

The meta-analytic results showed that MPH significantly improved math accuracy by 3.0% ($p = .001$) and math productivity by 7.8% ($p < .001$), see Figs. 2 and 3, respectively. Results from the four studies not qualifying for meta-analysis and reviewed in our qualitative synthesis (see Table E1) corroborate our meta-analytic findings.

Reading

For reading, meta-analytic results showed that improvements in accuracy with MPH were not significant (improved by 6.2%, $p = .089$), see Fig. 4. In contrast, MPH increased the number of reading items attempted ($d = .47$, $p < .001$), see Fig. 5.

Spelling

The results from our qualitative synthesis were inconclusive with only one out of three studies reporting significant improvements in spelling with MPH compared to placebo, see Table E1.

All effect sizes reflecting the effects of MPH on math and reading performance showed low heterogeneity, see Table 1.

Publication bias

Inspection of Egger funnel plots for publication bias indicated no asymmetry for math productivity, reading accuracy and reading attempted. Egger's test was significant for math accuracy indicating a risk for publication bias. Fail-safe n values indicated that the effects of MPH on math accuracy, math productivity and reading attempted were quite robust, whereas the effect of MPH on reading accuracy was not robust (Table 1). There was no significant relation between sample size and effect size for any of the dependent variables entered in the meta-analysis. Therefore, it is unlikely that publication bias meaningfully influenced results, with the exception of MPH effects on reading accuracy. Risk of bias of individual studies according to the Cochrane index for crossover trials was generally low, for details see supplementary Table E3.

Mediation and moderation

None of the potential mediators or moderators significantly interacted with the effects of MPH on math accuracy or productivity (all ps > .09). Supplementary Table E4 reports on the number of studies included in the meta-regression, Z values, 95% CI and p values.

Table 1 Meta-analytic results for the effects of MPH on academic performance

	Meta-analytic effect size				Heterogeneity			Publication bias				
	k	ES	95% CI	p	I^2, %	Q (df)	p	Fs N	p (EF)	95% CI (EF)	p (N)	p (RB)
Math accuracy	29	.030	.012, .048	.001	0	9.601 (28)	1.00	59	.04	.0036, 1.0021	.97	.29
Math productivity	17	.078	.043, .112	< .001	0	2.855 (16)	1.00	66	.20	− .2130, .9525	.91	.44
Reading accuracy	9	.062	− .009, .134	.089	0	1.389 (8)	.994	0	.35	− .7168, 1.7524	.77	.99
Reading attempted	5	.47	.302, .637	<.001	0	0.597 (4)	.963	34	.16	− .7057, 2.6173	.73	.79

MPH methylphenidate, *k* number of studies, *ES* effect size (risk difference for math accuracy, math productivity and reading accuracy, Cohen's *d* for reading attempted), *CI* confidence interval, *Fs N* fail-safe *n*, *p (EF)* p value of Egger funnel plot asymmetry, *p (N)* p value of the relation between sample size and effect size, *p (RB)*, p value of the relation between risk of bias and effect size
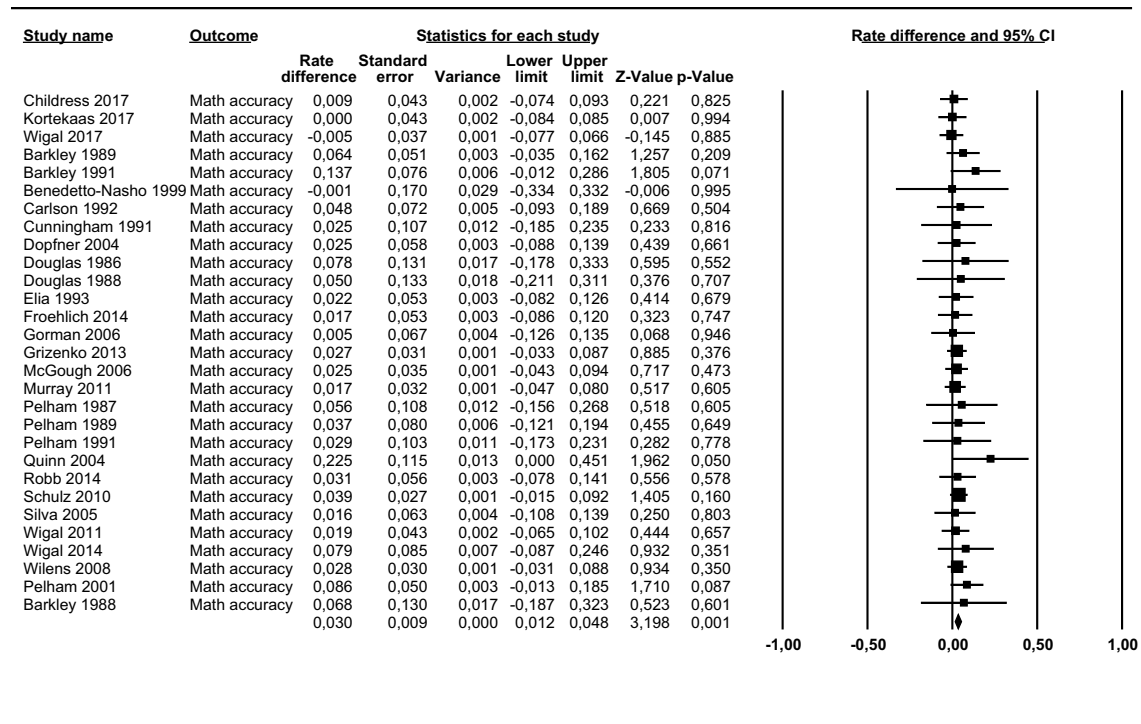
Study name	Outcome	Rate difference	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value	Rate difference and 95% CI
Childress 2017	Math accuracy	0,009	0,043	0,002	-0,074	0,093	0,221	0,825	
Kortekaas 2017	Math accuracy	0,000	0,043	0,002	-0,084	0,085	0,007	0,994	
Wigal 2017	Math accuracy	-0,005	0,037	0,001	-0,077	0,066	-0,145	0,885	
Barkley 1989	Math accuracy	0,064	0,051	0,003	-0,035	0,162	1,257	0,209	
Barkley 1991	Math accuracy	0,137	0,076	0,006	-0,012	0,286	1,805	0,071	
Benedetto-Nasho 1999	Math accuracy	-0,001	0,170	0,029	-0,334	0,332	-0,006	0,995	
Carlson 1992	Math accuracy	0,048	0,072	0,005	-0,093	0,189	0,669	0,504	
Cunningham 1991	Math accuracy	0,025	0,107	0,012	-0,185	0,235	0,233	0,816	
Dopfner 2004	Math accuracy	0,025	0,058	0,003	-0,088	0,139	0,439	0,661	
Douglas 1986	Math accuracy	0,078	0,131	0,017	-0,178	0,333	0,595	0,552	
Douglas 1988	Math accuracy	0,050	0,133	0,018	-0,211	0,311	0,376	0,707	
Elia 1993	Math accuracy	0,022	0,053	0,003	-0,082	0,126	0,414	0,679	
Froehlich 2014	Math accuracy	0,017	0,053	0,003	-0,086	0,120	0,323	0,747	
Gorman 2006	Math accuracy	0,005	0,067	0,004	-0,126	0,135	0,068	0,946	
Grizenko 2013	Math accuracy	0,027	0,031	0,001	-0,033	0,087	0,885	0,376	
McGough 2006	Math accuracy	0,025	0,035	0,001	-0,043	0,094	0,717	0,473	
Murray 2011	Math accuracy	0,017	0,032	0,001	-0,047	0,080	0,517	0,605	
Pelham 1987	Math accuracy	0,056	0,108	0,012	-0,156	0,268	0,518	0,605	
Pelham 1989	Math accuracy	0,037	0,080	0,006	-0,121	0,194	0,455	0,649	
Pelham 1991	Math accuracy	0,029	0,103	0,011	-0,173	0,231	0,282	0,778	
Quinn 2004	Math accuracy	0,225	0,115	0,013	0,000	0,451	1,962	0,050	
Robb 2014	Math accuracy	0,031	0,056	0,003	-0,078	0,141	0,556	0,578	
Schulz 2010	Math accuracy	0,039	0,027	0,001	-0,015	0,092	1,405	0,160	
Silva 2005	Math accuracy	0,016	0,063	0,004	-0,108	0,139	0,250	0,803	
Wigal 2011	Math accuracy	0,019	0,043	0,002	-0,065	0,102	0,444	0,657	
Wigal 2014	Math accuracy	0,079	0,085	0,007	-0,087	0,246	0,932	0,351	
Wilens 2008	Math accuracy	0,028	0,030	0,001	-0,031	0,088	0,934	0,350	
Pelham 2001	Math accuracy	0,086	0,050	0,003	-0,013	0,185	1,710	0,087	
Barkley 1988	Math accuracy	0,068	0,130	0,017	-0,187	0,323	0,523	0,601	
		0,030	0,009	0,000	0,012	0,048	3,198	0,001	

Fig. 2 Forest plot of the effects of MPH on math accuracy

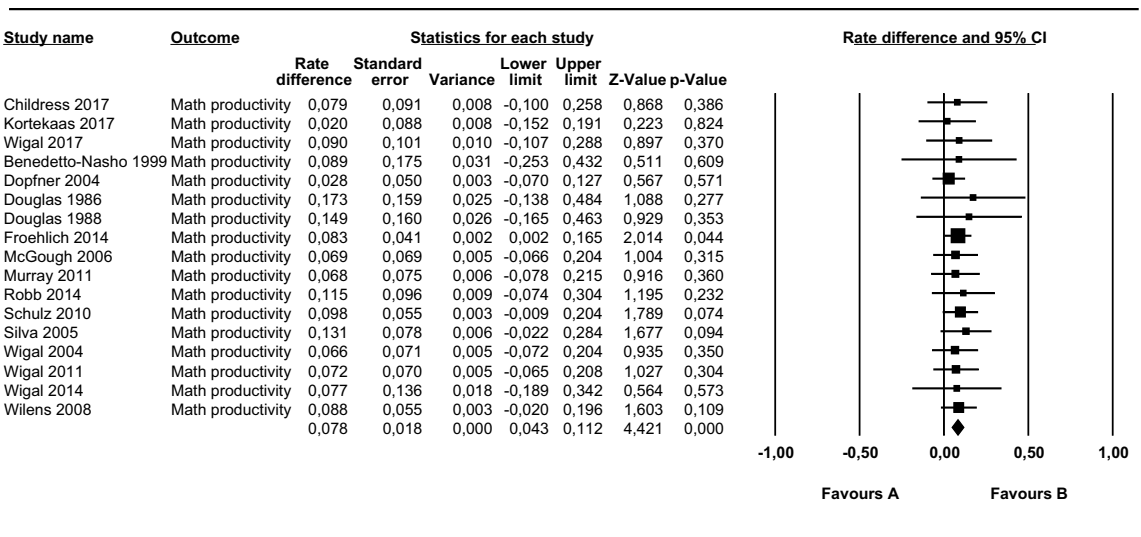| Study name	Outcome	Rate difference	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value	Rate difference and 95% CI
Childress 2017	Math productivity	0,079	0,091	0,008	-0,100	0,258	0,868	0,386	
Kortekaas 2017	Math productivity	0,020	0,088	0,008	-0,152	0,191	0,223	0,824	
Wigal 2017	Math productivity	0,090	0,101	0,010	-0,107	0,288	0,897	0,370	
Benedetto-Nasho 1999	Math productivity	0,089	0,175	0,031	-0,253	0,432	0,511	0,609	
Dopfner 2004	Math productivity	0,028	0,050	0,003	-0,070	0,127	0,567	0,571	
Douglas 1986	Math productivity	0,173	0,159	0,025	-0,138	0,484	1,088	0,277	
Douglas 1988	Math productivity	0,149	0,160	0,026	-0,165	0,463	0,929	0,353	
Froehlich 2014	Math productivity	0,083	0,041	0,002	0,002	0,165	2,014	0,044	
McGough 2006	Math productivity	0,069	0,069	0,005	-0,066	0,204	1,004	0,315	
Murray 2011	Math productivity	0,068	0,075	0,006	-0,078	0,215	0,916	0,360	
Robb 2014	Math productivity	0,115	0,096	0,009	-0,074	0,304	1,195	0,232	
Schulz 2010	Math productivity	0,098	0,055	0,003	-0,009	0,204	1,789	0,074	
Silva 2005	Math productivity	0,131	0,078	0,006	-0,022	0,284	1,677	0,094	
Wigal 2004	Math productivity	0,066	0,071	0,005	-0,072	0,204	0,935	0,350	
Wigal 2011	Math productivity	0,072	0,070	0,005	-0,065	0,208	1,027	0,304	
Wigal 2014	Math productivity	0,077	0,136	0,018	-0,189	0,342	0,564	0,573	
Wilens 2008	Math productivity	0,088	0,055	0,003	-0,020	0,196	1,603	0,109	
		0,078	0,018	0,000	0,043	0,112	4,421	0,000	

Favours A   Favours B

Fig. 3 Forest plot of the effects of MPH on math productivity

Discussion

The current meta-analysis and systematic review summarized more than three decades of research on the effects of MPH on academic performance in ADHD. Our analysis particularly focused on the question whether MPH improves academic accuracy or just academic productivity.

There were small to medium-sized, positive effects of MPH on math accuracy, math productivity and reading accuracy. Math accuracy increased by 3.0%, whereas math
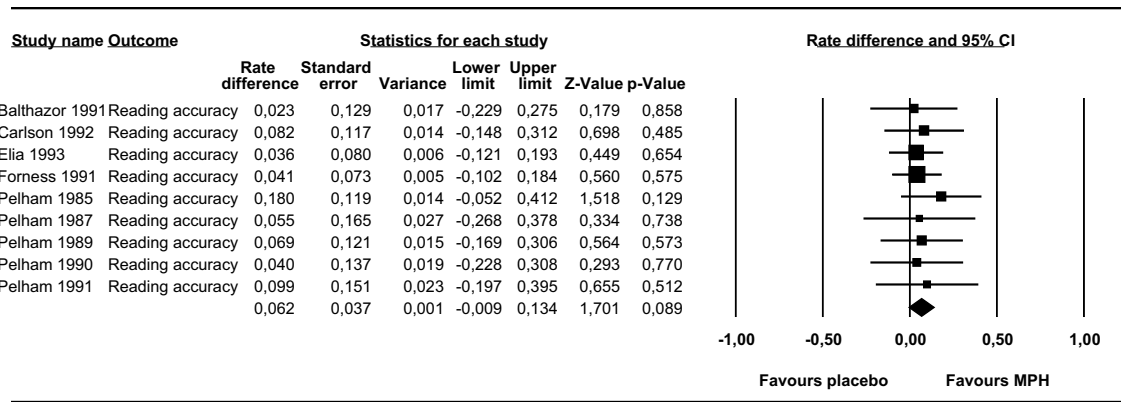
Study name	Outcome	Statistics for each study							Rate difference and 95% CI
		Rate difference	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value	
Balthazor 1991	Reading accuracy	0,023	0,129	0,017	-0,229	0,275	0,179	0,858	
Carlson 1992	Reading accuracy	0,082	0,117	0,014	-0,148	0,312	0,698	0,485	
Elia 1993	Reading accuracy	0,036	0,080	0,006	-0,121	0,193	0,449	0,654	
Forness 1991	Reading accuracy	0,041	0,073	0,005	-0,102	0,184	0,560	0,575	
Pelham 1985	Reading accuracy	0,180	0,119	0,014	-0,052	0,412	1,518	0,129	
Pelham 1987	Reading accuracy	0,055	0,165	0,027	-0,268	0,378	0,334	0,738	
Pelham 1989	Reading accuracy	0,069	0,121	0,015	-0,169	0,306	0,564	0,573	
Pelham 1990	Reading accuracy	0,040	0,137	0,019	-0,228	0,308	0,293	0,770	
Pelham 1991	Reading accuracy	0,099	0,151	0,023	-0,197	0,395	0,655	0,512	
		0,062	0,037	0,001	-0,009	0,134	1,701	0,089	

Favours placebo　　Favours MPH

Fig. 4 Forest plot of the effects of MPH on reading accuracy

Outcome	Statistics for each study							Std diff in means and 95% CI
	Std diff in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value	
Pelham 1990 Reading attempted	0,604	0,232	0,054	0,149	1,058	2,604	0,009	
Pelham 1987 Reading attempted	0,420	0,129	0,017	0,166	0,673	3,245	0,001	
Pelham 1989 Reading attempted	0,422	0,213	0,045	0,005	0,840	1,982	0,047	
Pelham 1991 Reading attempted	0,514	0,258	0,067	0,008	1,020	1,992	0,046	
Carlson 1992 Reading attempted	0,510	0,217	0,047	0,085	0,936	2,351	0,019	
	0,470	0,086	0,007	0,302	0,637	5,490	0,000	

Favours placebo　　Favours MPH

Fig. 5 Forest plot of the effects of MPH on reading number attempted

productivity increased by 7.8%. MPH did not improve reading accuracy, but did improve the number of items attempted in reading (medium effect). Results from our qualitative synthesis regarding math performance corroborated the findings from our meta-analyses. The qualitative synthesis of the studies reporting on spelling accuracy was inconclusive and more studies on this topic are needed. Our main results underline the importance of assessing a full range of outcome measures (accuracy and productivity) and different academic subjects when studying the effects of MPH on academic performance. Moreover, these results underline the contrast between the large symptom improvements obtained with MPH and the small- to medium-sized improvements in school performance—with improvements are restricted to certain academic subjects, and are small or absent for measures of accuracy.

However, it is important to realize that the short-term tests of academic performance used in the studies analyzed are sensitive to potential longer term benefits of MPH. Thus, it may be that although positive effects of MPH on academic accuracy are small or absent on the short term, MPH-related behavioral improvements and increased productivity may

result in long-term better school performance. Currently, evidence for long-term effects of MPH on academic performance is lacking (REF 14 Langberg) but cannot be ruled out because of methodological issues (i.e., the lack of long-term RCTs).

None of our mediators or moderators influenced MPH effects on math and reading accuracy or productivity. This may be because most variance is due to random error as indicated by the low I^2 values obtained in our meta-analyses [34]. However, as there is large uncertainty (large confidence intervals) in heterogeneity estimates such as I^2, we deemed our meta-regression relevant [47]. In particular, we hypothesized that teacher-rated symptom improvements would mediate MPH effects on academic performance, but the results from our meta-analysis did not confirm this hypothesis for MPH effects on math accuracy. Unfortunately, the number of studies reporting on teacher-rated symptom improvements and math productivity as well as reading was insufficient for meta-regression. Further, most studies reporting on on-task behavior and academic performance measures indicate simultaneous improvements on both with improvements in on-task behavior ranging from 2.9 to 12.0% [20, 24, 32,

48–55]. Unfortunately, the number of studies in the current study was too small to test the mediating effects of on-task behavior using meta-regression. Taken together, our results do not support a mediating role for classroom-expressed ADHD symptoms in the relationship between MPH and math accuracy, but this may be different for productivity measures as behavioral improvements in the classroom are generally seen as a prerequisite for academic improvements, especially academic productivity. Furthermore, cognitive improvements may be more relevant here than symptom improvements, as deficits of children with ADHD are apparent for those cognitive functions that are especially important for academic performance, e.g., attention, working memory and response inhibition [56–59]. Possibly MPH-related improvements in cognition play a large role in academic improvement, compared to behavioral improvements and these act through different pathways than those driving symptoms. This is consistent with a recent study by Coghill et al. [60] showing that while MPH improves both symptoms and some aspects of cognition these effects seem to be independent.

Demographic and disorder-related variables included in our analysis (age, gender, ADHD subtype and ADHD severity) did not moderate MPH efficacy on math performance either. The absence of a moderating effect of age and comorbid disorders is in line with the results of a recent meta-analysis on behavioral improvements with MPH [61]. In the meta-analysis by Storebø and colleagues, some evidence was found for a moderating effect of ADHD subtype and behavioral improvements, with highest MPH efficacy for the inattentive subtype.

Similarly, study characteristics also did not moderate efficacy of MPH on math performance, at least not for release system, trial duration, time of measurement, dosage and titration method. Thirteen studies used ER formulations, two studies reported on the effects of transdermal MPH and 19 studies reported on IR formulations. Results from our meta-regression suggest that ER formulations are equally effective as IR formulations in improving academic performance, which is in line with findings from individual studies comparing ER and IR formulations [52, 62, 63]. We found no effects of titration method (clinical titration prior to the trial or fixed dosages) which is also in line with findings from [61]. Possibly, to optimize the effects of MPH on academic outcome, titration should be based on academic outcomes instead of on symptom improvements. The absence of an effect of dose on academic performance is in line with the findings from Prasad et al. [13], who found no difference between studies comparing the effect of 0.3 mg/kg or 10 mg fixed dose to 0.6 mg/kg or 17.5–20 mg fixed dose on percentage seatwork completed. Also in line with this are the results from [61], who found no effects of dose on symptom improvements.

Strengths of the current review included a separate consideration of academic accuracy and productivity; including a distinction between academic subjects (math, reading and spelling) and the inclusion of randomized controlled trials only. There were, however, also some potential limitations. First, because we focused on effects with the optimal dose, we did not include dose–response analyses. Second, trial duration was generally short (between 1 and 7 days), limiting our conclusions to short-term effects of MPH on academic performance. Evidence for the longer term benefits of MPH on academic performance is lacking thus far [14]. For obvious ethical and practical reasons evidence for such effects is unlikely to be generated from placebo-controlled trials—and is, therefore, outside the scope of the current review. It remains possible that short-term effects of MPH on both behavior and academic performance (i.e., productivity) summarized here, may translate into longer term benefits. Furthermore, it should be investigated whether long-term benefits may be seen even where short-term effects are not evident.

On the basis of this review, we make a number of recommendations for future research. Some studies in the review relied on self-developed math or spelling test sheets. In the future, these should be replaced by validated tests (depending on the design of the study, with relevant norms). Further, researchers should always report both accuracy and productivity measures to allow for separate estimation of MPH effects on quantity and quality of academic performance. Moreover, by increasing the trial duration, more relevant measures like school grades can be included while still using randomized, placebo-controlled designs. Finally, more research on moderators and mediators of MPH efficacy is useful to isolate groups of patients who may benefit more or less from MPH and to reveal its mechanism of action. Although many studies have attempted to do so, measures of mediators and moderators are not uniform and generally not standardized, impeding (meta-analytic) aggregation of relevant results.

In summary, our results indicated that MPH results in robust improvements in the number of reading items attempted and small- to medium-sized improvements in math productivity and accuracy. Improvements in academic quality (accuracy) were small (3.0%) and limited to math. The effects of MPH on math accuracy were not mediated by teacher-rated ADHD symptom improvements, and MPH effects on math accuracy and productivity were not influenced by demographic variables, disorder-related variables or study characteristics. The discrepancy between the large behavioral improvements seen with MPH and these smaller and selective improvements in academic performance are important for treatment guidelines. As academic improvement is often one of the main treatment goals, parents and teachers should

be advised about the specificity and limited size of MPH effects on academic performance.

Compliance with ethical standards

Conflict of interest Anne Fleur Kortekaas-Rijlaarsdam MSc., Dr. Marjolein Luman has no conflicts of interest to declare. Prof. Dr. Jaap Oosterlaan (last 3 years): research funding (investigator initiated trial) from Shire Pharma. Prof. Dr. Edmund Sonuga-Barke (last 3 years): speaker fees, consultancy and conference support from Shire Pharma. Speaker fees from Janssen Cilag, consultancy from Neurotech solutions, Aarhus University, Copenhagen University and Berhanderling, Skolerne, Copenhagen, KU Leuven. Book royalties from OUP and Jessica Kingsley.

References

1. American Psychiatric Association (2013) Diagnostic and statistical manual of mental disorders, 5th edn. American Psychiatric Publishing, Arlington
2. Barry TD, Lyman RD, Klinger LG (2002) Academic underachievement and attention-deficit/hyperactivity disorder: the negative impact of symptom severity on school performance. J School Psychol 40:259–283
3. DuPaul GJ (2007) School-based interventions for students with attention deficit hyperactivity disorder: current status and future directions. School Psych Rev 36:183–194
4. Polderman TJC, Boomsma DI, Bartels M et al (2010) A systematic review of prospective studies on attention problems and academic achievement. Acta Psychiatr Scand 122:271–284. https://doi.org/10.1111/j.1600-0447.2010.01568.x
5. Frazier TW, Youngstrom EA, Glutting JJ, Watkins MW (2007) ADHD and achievement: meta-analysis of the child, adolescent, and adult literatures and a concomitant study with college students. J Learn Disabil 40:49–65. https://doi.org/10.1177/00222194070400010401
6. Wright N, Moldavsky M, Schneider J et al (2015) Practitioner review: pathways to care for ADHD—a systematic review of barriers and facilitators. J Child Psychol Psychiatry 56:598–617. https://doi.org/10.1111/jcpp.12398
7. MTA (1999) A 14-month randomized clinical trial of treatment strategies for attention-deficit/hyperactivity disorder. Arch Gen Psychiatry 56:1073–1086
8. Van der Oord S, Prins PJM, Oosterlaan J, Emmelkamp PMG (2008) Efficacy of methylphenidate, psychosocial treatments and their combination in school-aged children with ADHD: a meta-analysis. Clin Psychol Rev 28:783–800
9. Coghill DR, Seth S, Pedroso S et al (2013) Effects of methylphenidate on cognitive functions in children and adolescents with attention-deficit/hyperactivity disorder: evidence from a systematic review and a meta-analysis. Biol Psychiatry. https://doi.org/10.1016/j.biopsych.2013.10.005
10. Aman MG (1980) Psychotropic drugs and learning problems: a selective review. J Learn Disabil 13(2):87–97
11. Barkley RA (1977) A review of stimulant drug research with hyperactive children. J Child Psychol Psychiatry 18:137–165
12. Schachar R, Tannock R (1993) Hyperactivity and psychostimulant: of extended treatment studies a review. J Child Adolesc Psychopharmacol 3:81–97
13. Prasad V, Brogan E, Mulvaney C et al (2013) How effective are drug treatments for children with ADHD at improving on-task behaviour and academic achievement in the school classroom? A systematic review and meta-analysis. Eur Child Adolesc Psychiatry 22:203–216
14. Langberg JM, Becker SP (2012) Does long-term medication use improve the academic outcomes of youth with attention-deficit/hyperactivity disorder? Clin Child Fam Psychol Rev 15:215–233
15. Arnold LE, Hodgkins P, Kahle J et al (2015) Long-term outcomes of ADHD: academic achievement and performance. J Attent Disord. https://doi.org/10.1177/1087054714566076
16. Barkley RA, Varley CK, Nolan PA (1978) Do stimulant drugs improve the academic performance of hyperkinetic children? Clin Pediatr (Phila) 17:85–92
17. Schachar R, Tannock R (1993) Childhood hyperactivity and psychostimulants: a review of extended treatment studies. J Child Adolesc Psychopharmacol 3:81–97
18. Murray DW, Childress A, Giblin J et al (2011) Effects of OROS methylphenidate on academic, behavioral, and cognitive tasks in children 9 to 12 years of age with attention-deficit/hyperactivity disorder. Clin Pediatr (Phila) 50:308–320
19. Kortekaas-Rijlaarsdam AF, Luman M, Sonuga-Barke E et al (2017) Short-term effects of methylphenidate on math productivity in children with attention-deficit/hyperactivity disorder are mediated by symptom improvements: evidence from a placebo-controlled trial. J Clin Psychopharmacol 37:210–219
20. Froehlich TE, Antonini TN, Brinkman WB et al (2014) Mediators of methylphenidate effects on math performance in children with attention-deficit hyperactivity disorder. J Dev Behav Pediatr 35:100–107
21. Moher D, Liberati A, Tetzlaff J, Altman DG (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 6:e1000097. https://doi.org/10.1371/journal.pmed.1000097
22. Thiruchelvam D, Charach A, Schachar RJ (2001) Moderators and mediators of long-term adherence to stimulant treatment in children with ADHD. J Am Acad Child Adolesc Psychiatry 40:922–928
23. GetData GRaph Digitizer [Computer software] (2013). Retrieved from http://getdata-graphdigitizer.com
24. Pelham WE, Gnagy EM, Burrows-Maclean L et al (2001) Once-a-day Concerta methylphenidate versus three-times-daily methylphenidate in laboratory and natural settings. Pediatrics 107:E105. https://doi.org/10.1542/peds.107.6.e105
25. Williamson D, Murray DW, Damaraju CV et al (2014) Methylphenidate in children With ADHD with or without learning disability. J Atten Disord 18:95–104
26. Armstrong RB, Damaraju CV, Ascher S et al (2012) Time course of treatment effect of OROS(registered trademark) methylphenidate in children with ADHD. J Atten Disord 16:697–705
27. Forness SR, Swanson JM, Cantwell DP et al (1992) Stimulant medication and reading performance: follow-up on sustained dose in ADHD boys with and without conduct disorders. J Learn Disabil 25:115–123
28. Forness SR, Swanson JM, Cantwell D et al (1992) Response to stimulant medication across six measures of school-related performance in children with ADHD and disruptive behavior. Behav Disord 18:42–53
29. Wigal SB, Gupta S, Heverin E, Starr HL (2011) Pharmacokinetics and therapeutic effect of OROS(registered trademark) methylphenidate under different breakfast conditions in children with

attention-deficit/hyperactivity disorder. J Child Adolesc Psychopharmacol 21:255–263

30. Forness SR, Cantwell DP, Swanson JM et al (1991) Differential effects of stimulant medication on reading performance of boys with hyperactivity with and without conduct disorder. J Learn Disabil 24:304–310

31. Douglas VI, Barr RG, O'Neill ME, Britton BG (1986) Short term effects of methylphenidate on the cognitive, learning and academic performance of children with attention deficit disorder in the laboratory and the classroom. J Child Psychol Psychiatry Allied Discip 27:191–211

32. Pelham WE, Bender ME, Caddell J (1985) Methylphenidate and children with attention deficit disorder. Dose effects on classroom academic and social behavior. Arch Gen Psychiatry 42:948–952

33. Bental B, Tirosh E (2008) The effects of methylphenidate on word decoding accuracy in boys with attention-deficit/hyperactivity disorder. J Clin Psychopharmacol 28:89–92

34. Borenstein M, Hedges L, Higgins J, Rothstein H (2009) Introduction to meta-analysis. Wiley, Chichester

35. Loney J, Milich R (1982) Hyperactivity, inattention, and aggression in clinical practice. In: Wolraich M, Routh D (eds) Advances in developmental and behavioral pediatrics, vol 3. JAI Press, Greenwich, CT, pp 113–147

36. Swanson J, Schuck S, Mann M, et al (1994) Categorical and dimensional definitions and evaluations of symptoms of ADHD. The SNAP and SWAN ratings scales [Draft]. pp 1–20. http://www.adhd.net/SNAP_SWAN.pdf. Accessed 14 Apr 2014

37. Hay DA, Bennett KS, Levy F et al (2007) A twin study of attention-deficit/hyperactivity disorder dimensions rated by the strengths and weaknesses of ADHD-symptoms and normal-behavior (SWAN) scale. Biol Psychiatry 61:700–705. https://doi.org/10.1016/j.biopsych.2006.04.040

38. Wigal SB, Gupta S, Guinta D, Swanson JM (1998) Reliability and validity of the SKAMP rating scale in a laboratory school setting. Psychopharmacol Bull 34:47–53

39. Conners CK (1997) Conners' global index-teacher. Multihealth Systems Inc., North Tonawanda, NY

40. IBM (2012) IBM SPSS statistics for windows, version 21.0

41. Borenstein M (2014) Comprehensive meta-analysis V3.0

42. Twisk JWR (2010) Introduction in applied biostatistics (Inleiding in de toegepaste biostatistiek). Reed Business, Amsterdam

43. Higgins JPT, Thompson SG, Deeks JJ, Altman DG (2003) Measuring inconsistency in meta-analyses. BMJ Br Med J 327:557–560. https://doi.org/10.1136/bmj.327.7414.557

44. Rosenthal R (1995) Writing meta-analytic reviews. Psychol Bull 118:183–192. https://doi.org/10.1037/0033-2909.118.2.183

45. Egger M, Smith GD, Schneider M, Minder C (1997) Bias in meta-analysis detected by a simple, graphical test. BMJ 315:629–634

46. Higgins J, Green S (2011) Cochrane handbook for systematic reviews of interventions version 5.1.0 [updated March 2011]

47. Ioannidis JPA, Patsopoulos NA, Evangelou E (2007) Uncertainty in heterogeneity estimates in meta-analyses. BMJ 335:914–916. https://doi.org/10.1136/bmj.39343.408449.80

48. Gorman EB, Klorman R, Thatcher JEBA (2006) Effects of methylphenidate on subtypes of attention-deficit/hyperactivity disorder. J Am Acad Child Adolesc Psychiatry 45:808–816. https://doi.org/10.1097/01.chi.0000214191.57993.dd

49. Carlson CL, Pelham WE Jr, Milich R, Dixon J (1992) Single and combined effects of methylphenidate and behavior therapy on the classroom performance of children with attention-deficit hyperactivity disorder. J Abnorm Child Psychol 20:213–232

50. Cunningham CE, Siegel LS, Offord DR (1991) A dose-response analysis of the effects of methylphenidate on the peer interactions and simulated classroom performance of ADD children with and without conduct problems. J Child Psychol Psychiatry Allied Discip 32:439–452

51. Pelham WE, Vodde-Hamilton M, Murphy DA et al (1991) The effects of methylphenidate on ADHD adolescents in recreational, peer group, and classroom settings. J Clin Child Psychol 20:293–300

52. Pelham WE Jr, Greenslade KE, Vodde-Hamilton M et al (1990) Relative efficacy of long-acting stimulants on children with attention deficit-hyperactivity disorder: a comparison of standard methylphenidate, sustained-release methylphenidate, sustained-release dextroamphetamine, and pemoline. Pediatrics 86:226–237

53. Pelham WE Jr, Walker JL, Sturges J, Hoza J (1989) Comparative effects of methylphenidate on ADD girls and ADD boys. J Am Acad Child Adolesc Psychiatry 28:773–776

54. Pelham WE Jr, Sturges J, Hoza J et al (1987) Sustained release and standard methylphenidate effects on cognitive and social behavior in children with attention deficit disorder. Pediatrics 80:491–501

55. Barkley RA, Fischer M, Newby RF, Breen MJ (1988) Development of a multimethod clinical protocol for assessing stimulant drug response in children with attention deficit disorder. J Clin Child Psychol 17:14–24

56. Mayes SD, Calhoun SL (2007) Learning, attention, writing, and processing speed in typical children and children with ADHD, autism, anxiety, depression, and oppositional-defiant disorder. Child Neuropsychol 13:469–493. https://doi.org/10.1080/09297040601112773

57. Preston AS, Heaton SC, McCann SJ et al (2009) The role of multidimensional attentional abilities in academic skills of children with ADHD. J Learn Disabil 42:240–249. https://doi.org/10.1177/0022219408331042

58. Thorell LB (2007) Do delay aversion and executive function deficits make distinct contributions to the functional impact of ADHD symptoms? A study of early academic skill deficits. J Child Psychol Psychiatry Allied Discip 48:1061–1070. https://doi.org/10.1111/j.1469-7610.2007.01777.x

59. Biederman J, Monuteaux MC, Doyle AE et al (2004) Impact of executive function deficits and attention-deficit/hyperactivity disorder (ADHD) on academic outcomes in children. J Consult Clin Psychol 72:757–766. https://doi.org/10.1037/0022-006X.72.5.757

60. Coghill DR, Hayward D, Rhodes SM et al (2014) A longitudinal examination of neuropsychological and clinical functioning in boys with attention deficit hyperactivity disorder (ADHD): improvements in executive functioning do not explain clinical improvement. Psychol Med 44:1087–1099. https://doi.org/10.1017/S0033291713001761

61. Storebø O, Krogh H, Ramstad E et al (2015) Methylphenidate for attention-deficit/hyperactivity disorder in children and adolescents: Cochrane systematic review with meta-analyses and trial sequential analyses of randomised clinical trials. Br Med J 351:1–14. https://doi.org/10.1136/bmj.h5203

62. Dopfner M, Gerber WD, Banaschewski T et al (2004) Comparative efficacy of once-a-day extended-release methylphenidate, two-times-daily immediate-release methylphenidate, and placebo in a laboratory school setting. Eur Child Adolesc Psychiatry, Suppl 13:I/93–I/101

63. Swanson JM, Wigal SB, Wigal T et al (2004) A comparison of once-daily extended-release methylphenidate formulations in children with attention-deficit/hyperactivity disorder in the laboratory school (the Comacs Study). Pediatrics 113:e206–e216